

**KOMPARASI DATA MINING UNTUK MEMPREDIKSI MINAT KLIEN
DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN
(Studi pada PT. AJB Bumiputera 1912 Karanganyar)**



**Disusun sebagai salah satu syarat menyelesaikan Program Studi Strata I pada Jurusan
Informatika Fakultas Komunikasi dan Informatika**

Oleh:

HERIA YUNITA PUTRI

L 200 140 063

**PROGRAM STUDI INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMTIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA
2018**

HALAMAN PERSETUJUAN

**KOMPARASI DATA MINING UNTUK MEMPREDIKSI MINAT KLIEN
DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN
(Studi pada PT. AJB Bumiputera 1912 Karanganyar)**

PUBLIKASI ILMIAH

oleh:

HERIA YUNITA PUTRI

L 200 140 063

Telah diperiksa dan disetujui untuk diuji oleh:

Dosen Pembimbing



Umi Fadilah, S.T. M.Eng.
NIP. 197803222005012002

HALAMAN PENGESAHAN

**KOMPARASI DATA MINING UNTUK MEMPREDIKSI MINAT KLIEN
DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN
(Studi pada PT. AJB Bumiputera 1912 Karanganyar)**

OLEH

HERIA YUNITA PUTRI

L 200 140 063

Telah dipertahankan di depan Dewan Penguji
Fakultas ~~Komunikasi dan Informatika~~
Universitas Muhammadiyah Surakarta
Pada hari ~~Jumat~~ 19-01-2018
dan dinyatakan telah memenuhi syarat

Dewan Penguji:

1. Umi Fadlilah, S.T., M.Eng.
(Ketua Dewan Penguji)
2. Dr. Heru Supriyono, M.Sc.
(Anggota I Dewan Penguji)
3. Aris Rakhmadi, S.T., M.Eng.
(Anggota II Dewan Penguji)

(.....)
(.....)
(.....)

Publikasi ilmiah ini telah diterima sebagai salah satu persyaratan

Untuk memperoleh gelar sarjana

Tanggal 31-01-2018

Mengetahui,



Dekan
Fakultas Komunikasi dan Informatika

Nurgiyatna, M.Sc., Ph.D.
NIK : 881



Ketua Program Studi
Informatika

Dr. Heru Supriyono, M.Sc.
NIK : 970

PERNYATAAN

Dengan ini saya **menyatakan** bahwa dalam naskah publikasi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam naskah dan disebutkan dalam daftar pustaka.

Apabila kelak terbukti ada ketidakbenaran dalam pernyataan saya di atas, maka akan saya pertanggungjawabkan sepenuhnya.

Surakarta, 12 Januari 2018

Penulis



HERIA YUNITA PUTRI

L 200 140 063



**UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI INFORMATIKA**

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id>. Email: informatika@ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

No Surat 021/A.3-11-3/INF-FK/1/2018

Assalamu'alaikum Wr. Wb

Biro Skripsi Program Studi Informatika menerangkan bahwa :

Nama : Heria Yunita Putri
NIM : **L 200 140 063**
Judul : Komparasi Data Mining Untuk Memprediksi Minat Klien Dalam
Memilih Produk Asuransi Pendidikan (Studi pada PT. AJB Bumiputera
1912 Karanganyar)
Program Studi : Informatika
Status : **Lulus**

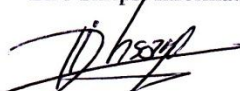
Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Skripsi, dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 26 Januari 2018

Biro Skripsi Informatika


Ihsan Cahyo Utomo, S.Kom., M.Kom.



UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI INFORMATIKA

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id>. Email: informatika@ums.ac.id

Feedback Studio - Google Chrome
Secure | https://extumitn.com/app/carta/en_us/?o=506057539&u=10575500800&s=1&lang=en_us

feedback studio KOMPARASI DATA MINING UNTUK MEMPREDEKSI MINAT KLIEN DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN (SI) -- /0 < > ?

KOMPARASI DATA MINING UNTUK MEMPREDEKSI MINAT KLIEN DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN
(Studi pada PT. AJB Bumiputera 1912 Karanganyar)
Heria Yunita Putri, Umi Fadillah, S.T., M.Eng.

Abstrak

Asuransi pendidikan PT. AJB Bumiputera 1912 terdapat dua jenis produk yaitu mitra beasiswa dan mitra cerdas. *Data mining* merupakan solusi gudang data masyarakat yang berguna bagi perusahaan dalam menemukan informasi penting. Penelitian ini bertujuan membantu perusahaan dalam menawarkan produk agar mutu produk tetap terjaga dengan menganalisis data mentah perusahaan asuransi untuk mengetahui faktor mana yang paling berpengaruh terhadap minat klien pada produk asuransi pendidikan menggunakan perbandingan 3 metode algoritma *data mining*. Metode algoritma *data mining* yang digunakan yaitu metode algoritma *Information Gain*, *Gain Ratio* dan *Rule Induction*. Dalam pengolahan datanya dibantu dengan aplikasi *rapid miner* untuk membantu mengetahui hasil *accuracy*, *precision*, dan *recall*. Alat yang digunakan yaitu macam asuransi (X1), cara bayar (X2), uang peranggungan (X3), premi (X4), dan minat (Y). Hasil dari implementasi penelitian ini memperoleh nilai *accuracy*, *precision*, dan *recall* pada metode *information gain* dan

Match Overview

16%

Rank	Source	Percentage
1	ejournal.ums.ac.id	8%
2	Submitted to Universitas...	2%
3	Mohamed K. Elhadad...	1%
4	artikelasuransi@waterb...	1%
5	media.neliti.com	1%
6	jurnal.ums.ac.id	1%
7	Submitted to Universitas...	1%

Page: 1 of 14 Word Count: 3071

16:45 AM 12/06/2019

KOMPARASI DATA MINING UNTUK MEMPREDIKSI MINAT KLIEN DALAM MEMILIH PRODUK ASURANSI PENDIDIKAN (Studi pada PT. AJB Bumiputera 1912 Karanganyar)

Abstrak

Asuransi pendidikan PT. AJB Bumiputera 1912 terdapat dua jenis produk yaitu mitra beasiswa dan mitra cerdas. *Data mining* merupakan solusi gudang data masyarakat yang berguna bagi perusahaan dalam menemukan informasi penting. Penelitian ini bertujuan membantu perusahaan dalam menawarkan produk agar mutu produk tetap terjaga dengan menganalisis data mentah perusahaan asuransi untuk mengetahui faktor manakah yang paling berpengaruh terhadap minat klien pada produk asuransi pendidikan menggunakan perbandingan 3 metode algoritma *data mining*. Metode algoritma *data mining* yang digunakan yaitu metode algoritma *Information Gain*, *Gain Ratio* dan *Rule Induction*. Dalam pengolahan datanya dibantu dengan aplikasi *rapid miner* untuk membantu mengetahui hasil *accuracy*, *precisison*, dan *recall*. Atribut yang digunakan yaitu macam asuransi (X1), cara bayar (X2), uang pertanggungan (X3), premi (X4), dan minat (Y). Hasil dari implementasi penelitian ini memperoleh nilai *accuracy*, *precision*, dan *recall* pada metode *information gain* dan *gain ratio* memiliki nilai tinggi yang sama dibandingkan dengan metode *rule induction* dengan nilai *accuracy* sebesar 85,69 %, nilai *precision* sebesar 99,18 %, dan nilai *recall* sebesar 54,50 %. Metode *information gain* dan *gain ratio* dalam penelitian ini menjadi metode yang sama unggul. Variabel uang pertanggungan adalah variabel paling berpengaruh pada penelitian minat klien dalam memilih produk asuransi pendidikan. Jadi, perusahaan dalam menawarkan asuransi pendidikan harus memperhatikan jumlah uang petanggungan terhadap premi yang dibayarkan agar tetap menjaga mutu produk asuransi mitra cerdas maupun mitra beasiswa.

Kata Kunci: asuransi, *data mining*, *gain ratio*, *information gain*, *rule induction*.

Abstract

Education Insurance PT AJB Bumiputera 1912 is classified into 2 kind of products namely mitra beasiswa and mitra cerdas. Data mining is society data warehouse solution which is beneficial for the company to find the important information. The study aims at helping the company to offer the products so that the product quality is maintained through analysing the raw data of the insurance company to encounter the influenced factor toward clients' interest on education insurance products by using comparison of three methods algorithm of data mining. Algorithm of data mining used is algorithm method of information gain, gain ratio and rule induction. On data analysis is assisted by rapid manner application to know the accuracy result, precision result and recall result. The used attributes are kinds of insurances (X1), payment method (X2), sum insured (X3), premium (X4) and interest (Y). The result of the implemented study gains the percentage of accuracy, precision and recall on information gain and gain ratio method have same high percentages compared with rule induction method with its percentage of accuracy is 85, 69% while precision is 99, 18% and recall is 54, 50%. Information gain and gain ratio methods on this study become equally superior. The sum insured variable is the most influential variable in the terms of clients' interest in choosing the education insurance products. Therefore, the offered education insurance of the company should

pay more attention to the amount of money insured towards the premium that is paid to keep the products quality both of mitra cerdas insurance or mitra beasiswa.

Keywords: *data mining, gain ratio, information gain, insurance, rule induction.*

1. PENDAHULUAN

Pendidikan adalah hal penting bagi kelangsungan hidup ke depan. Untuk memenuhi dana pendidikan tidaklah murah, orang tua harus menyiapkan dana pendidikan yang cukup tinggi dengan seiringnya dana yang terus naik tahun demi tahun. PT. Asuransi Jiwa Bumiputera 1912 adalah perusahaan asuransi terkemuka dan perusahaan asuransi jiwa nasional terbesar di Indonesia yang sudah berdiri sejak 105 tahun yang lalu. Asuransi pendidikan pada PT. AJB Bumiputera 1912 terdapat 2 jenis yaitu mitra beasiswa dan mitra cerdas. Mitra beasiswa memberikan perlindungan anak dalam biaya pendidikan yang dimulai dari taman kanak-kanak hingga perguruan tinggi. Sedangkan mitra cerdas dirancang untuk biaya pendidikan yang meningkat sejalan dengan hasil investasi.

Perusahaan asuransi pasti menyimpan data nasabah mereka dengan jumlah besar. Data mentah nasabah dapat diolah menjadi suatu keputusan yang bermanfaat bagi perusahaan. Untuk mengolah data mentah nasabah membutuhkan waktu yang tidak singkat karena keterbatasan waktu. *Data mining* merupakan solusi gudang data masyarakat yang berguna bagi perusahaan dalam menemukan informasi penting. (Anita, 2015).

Lhorend Mutiara Pratiwi, Diana, dan Eka Puji Agustin (2017), dalam penelitiannya yang berjudul Penerapan *K-means Clustering* untuk Memprediksi Minat Nasabah pada PT. Asuransi Jiwa Bersama 1912 Bumiputera Prabumulih, menyebutkan bahwa penelitian ini menggunakan data nasabah yang diolah menggunakan aplikasi weka dengan atribut macam asuransi, pekerjaan, pendapatan, dan alamat sebagai proses *data mining*. Penelitian tersebut menunjukkan bahwa nasabah yang berprofesi sebagai petani dengan jenis asuransi mitra beasiswa di wilayah gelombang memiliki minat tinggi dalam menjadi nasabah asuransi.

Lis Mahani (2017), dalam penelitiannya yang berjudul Perbandingan 3 Metode Dalam Data Mining untuk Penentuan Kadar Kolesterol di RSUD Dr. Moewardi Surakarta menyebutkan bahwa penelitian ini menggunakan metode perbandingan *data mining decision tree, naive bayes*, dan *rule induction*. Pengumpulan data penelitian ini menggunakan data pasien pemeriksaan kolesterol tahun 2015 yang berjumlah 9000 pasien tetapi peneliti menggunakan rumus metode *slovin* sehingga jumlah sampel yang diambil yaitu 383 pasien. Alat yang digunakan untuk mengolah data tersebut menggunakan *rapid miner 5*. Penelitian ini bertujuan ingin menentukan faktor yang paling berpengaruh dalam penentuan kadar kolesterol dan metode *data mining* manakah yang paling baik dalam penelitian tersebut.

Pradnya Kumbhar dan Manisha Mali (2013), dalam penelitiannya yang berjudul *A Survey on Feature Selection Techniques and Classification Algorithms for Efficient Text Classification*, menyebutkan bahwa penelitian ini mengkaji klasifikasi teks, beberapa pendekatan klasifikasi teks, metode pemilihan fitur dan aplikasi teks klasifikasi menggunakan banyak algoritma yaitu algoritma *naive bayes*, *decision tree induction*, *k-nearest neighbor*, *support vector machines* dan *neural network*. Tujuan dari penelitian ini yaitu peneliti ingin memberi tahu jenis klasifikasi teks dan langkah pengklasifikasian data yang nantinya bisa digunakan dalam pembuatan *data mining*.

Berdasarkan masalah diatas, belum ada yang meneliti tentang minat klien secara khusus untuk produk asuransi pendidikan dengan menggunakan perbandingan 3 metode *data mining information gain*, *gain ratio*, dan *rule induction*. Penelitian ini untuk mengetahui metode manakah yang paling akurat dalam menentukan minat klien terhadap produk asuransi pendidikan PT. AJB Bumiputera 1912 Karanganyar.

2. METODE

2.1 Penentuan Atribut dan Variabel

Untuk menentukan variabel terdapat 2 jenis variabel yaitu variabel independen (X) dan variabel dependen (Y). Variabel independen (X) adalah variabel yang nilainya tidak tergantung pada nilai variabel lainnya, sedangkan variabel dependen (Y) adalah variabel yang nilainya terikat pada nilai variabel lainnya. Jadi, variabel dan atribut yang digunakan dalam proses *data mining* ini adalah variabel yang sesuai dengan penelitian dan dapat diolah pada aplikasi *rapid miner 5*, yaitu seperti pada tabel 1 :

Tabel 1. Variabel dan Atribut

Variabel	Atribut
X1	Macam asuransi
X2	Cara bayar
X3	Uang pertanggungan
X4	Premi
Y	Minat

2.2 Pengumpulan Data

Penelitian ini menggunakan data nasabah PT. AJB Bumiputera 1912 Karanganyar dengan jumlah data awal 1670. Setelah dilakukan penyaringan sesuai dengan atribut yang dipilih, data yang digunakan dalam penelitian ini berjumlah 1069 dengan pembagian 2/3 sebagai data *training* sebanyak 713 data dan 1/3 sebagai data *testing* sebanyak 356 data.

2.3 Information Gain

Khumbur, dkk (2013) menyatakan *information gain* digunakan untuk mengukur jumlah informasi dalam bit tentang prediksi kelas sesuai distribusi, secara konkret mengukur ketidakpastian kelas yang terkait secara acak. Persyaratan seleksi atribut *information gain* yaitu memiliki nilai *gain* tertinggi. Rumus *information gain* terdapat pada **persamaan 1 dan 2**.

Persamaan 1 Rumus *Entropy* :

$$Entropy = \sum_{i=1}^n -p_i \log_2 p_i \quad (1)$$

Persamaan 2 Rumus *Information Gain* :

$$Gain(A) = entropy(S) - \sum_{i=1}^m \frac{S_k}{S} entropy(S_k) \quad (2)$$

Keterangan :

n = jumlah kelas

Pi = probabilitas S termasuk kelas i

Sk = subset dari S

2.4 Gain Ratio

Farahdiba, dkk (2016) menyatakan nilai *gain ratio* tertinggi ditempatkan menjadi node pertama pada decision tree. Mencari nilai *gain ratio* terlebih dahulu mencari nilai *split information*, terdapat pada **persamaan 3**.

$$split\ information\ (Y, A) = \sum_{i=1}^c (|Y_i| / |Y|) \log_2 (|Y_i| / |Y|) \quad (3)$$

Jadi, nilai *gain ratio* digunakan rumus pada **persamaan 4**.

$$gain\ ratio\ (S, A) = \frac{information\ gain\ (S, A)}{split\ information\ (S, A)} \quad (4)$$

2.5 Rule Induction

Ochtario (2013) menyatakan *rule induction* digunakan mencari sekumpulan item yang sering muncul bersamaan. Nugroho (2016) menyatakan rumus *rule induction* terdapat pada **persamaan 5, 6, 7**.

- a. Untuk mendapatkan nilai *support* untuk sebuah item

$$Support(A) = \frac{jumlah\ transaksi\ yang\ mengandung\ item\ A}{Total\ transaksi} \quad (5)$$

- b. Untuk mencari nilai *support* dari 2-item

$$Support\ (A, B) = P\ (A \cap B)$$

$$P(A \cap B) = \frac{jumlah\ transaksi\ yang\ mengandung\ A\ dan\ B}{Total\ transaksi} \quad (6)$$

c. Mencari nilai *confidence*

$$\text{Confidence } (A \rightarrow B) = P(B | A)$$

$$P(B | A) = \frac{\text{jumlah transaksi yang mengandung A dan B}}{\text{jumlah transaksi yang mengandung item A}} \quad (7)$$

3. HASIL DAN PEMBAHASAN

Data asli nasabah PT. AJB Bumiputera 1912 Karanganyar, ditunjukkan pada gambar 1 :

NO	STATUS	NOPOL	NAMA PP	MA	CB	JW	MULAS	U.PERTANGGUNGAN	PREMI	BYR AKHIR	KODE AGEN	ALAMAT
1	P	2000139162	RUKINI	WP	6	13	05/00	12.500.000,00	1.452.000	04/13	00/NMH	NGABLAK RT 05 RW 03 PAPAHAN TAS
2	P	98025308	SUMARNI	D6	5	0	02/98	3.000.000,00	2.042.004	01/15	00/KRA	DARASAN RT 02 RW 05 DALEMAN NG
3	P	2004450194	WIWIN SRI HANDAYANI	PM	5	0	11/04	5.000.000,00	1.000.000	10/19	00/KRA	PERUM PATAL NO.02 RT 1 RW 6 TAMI
4	P	2006027217	RINA HARSINIWATI SPD	MB	2	17	02/06	5.500.000,00	158.587	10/17	01/NMG	SAMBEREJO RT 01 RW 06 GEBYOG MC
5	P	2004097961	SAMI	DP	2	15	03/04	10.000.000,00	114.400	11/17	01/NDD	NGEPOS RT 02 RW 04 GAYAMDOMPO
6	P	2004097164	SOEMARYONO, BSC	D6	3	16	03/04	10.000.000,00	449.922	02/18	01/NPP	PENDEM KULON RT 02 RW 01 SURUH
7	P	2006046743	SRI HARTONO	DR	2	15	03/06	25.000.000,00	398.450	11/17	01/NMG	NGIRI RT 02 RW 03 NGEMPLAK KARAN
8	P	2006051652	HARTATIK	MB	2	16	03/06	5.000.000,00	153.530	11/17	01/NMG	GUDANG RT 01 RW 07 PASEBAN JUMI

Gambar 1. Sampel data asli nasabah PT. AJB Bumiputera 1912 Karanganyar

Dalam proses *data mining* menggunakan *rapid miner 5*, data dikelompokkan terlebih dahulu. Tipe data yang digunakan adalah tipe data *polynomial* dan tipe data *binominal*. Tipe data *binominal* yaitu tipe data yang terdiri dari 2 tipe, contohnya “Ya”, “Tidak”. Tipe data *polynomial* yaitu tipe data yang terdiri lebih dari 3 tipe, contohnya “Tinggi”, “Rendah”, “Sedang”. Berdasarkan data tersebut diambil atribut yang diperlukan dan dilakukan konversi data yang telah disetujui oleh perusahaan sehingga dapat diolah dan terbaca menggunakan *rapid miner 5*, ditunjukkan pada gambar 2 :

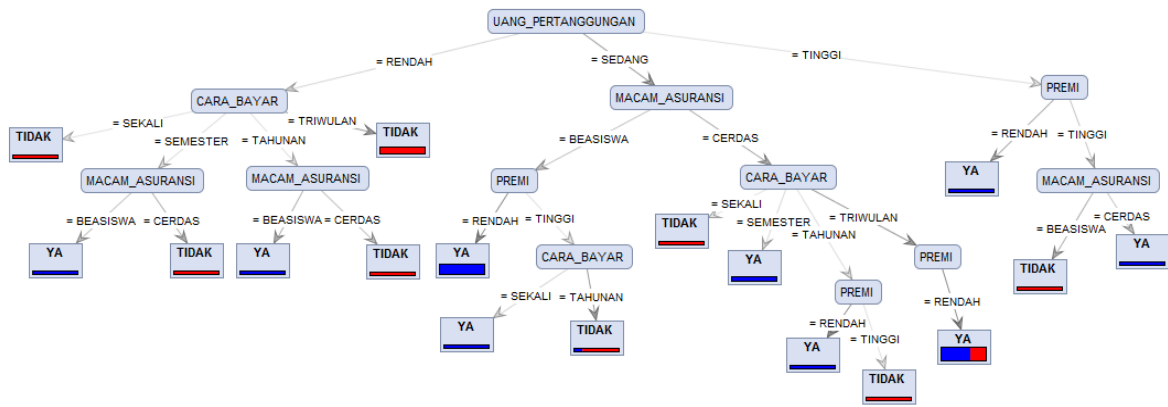
1	MACAM_ASURANSI	CARA_BAYAR	UANG_PERTANGGUNGAN	PREMI	MINAT
2	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
3	CERDAS	SEMESTER	SEDANG	RENDAH	YA
4	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
5	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
6	CERDAS	TRIWULAN	SEDANG	RENDAH	YA
7	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
8	BEASISWA	TRIWULAN	RENDAH	RENDAH	TIDAK
9	CERDAS	TRIWULAN	SEDANG	RENDAH	YA
10	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
11	CERDAS	TRIWULAN	SEDANG	RENDAH	TIDAK
12	CERDAS	TAHUNAN	SEDANG	RENDAH	YA
13	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
14	CERDAS	TRIWULAN	TINGGI	RENDAH	YA
15	BEASISWA	SEMESTER	SEDANG	RENDAH	YA
16	BEASISWA	SEMESTER	SEDANG	RENDAH	YA
17	CERDAS	SEKALI	RENDAH	TINGGI	TIDAK
18	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA
19	BEASISWA	TRIWULAN	SEDANG	RENDAH	YA

ATRIBUT	ATURAN KONVERSI	
Macam asuransi	D6	Cerdas
	MB	Beasiswa
Cara bayar	2	Triwulan
	3	Semester
	4	Tahunan
	5	Sekali
Uang Pertanggungan	< 5.000.000	rendah
	5.000.000 - 10.000.000	sedang
	> 10.000.000	tinggi
Premi	< 1.000.000	rendah
	> 1.000.000	tinggi

Gambar 2. Sampel data nasabah yang sudah dipilah dan konversi data

3.1 Implementasi *Information Gain* dan *Gain Ratio* pada *Rapid Miner 5*

Hasil proses *information gain* dan *gain ratio* untuk memprediksi minat klien dalam memilih produk asuransi menggunakan aplikasi *rapid miner 5* ditunjukkan pada gambar 3 :

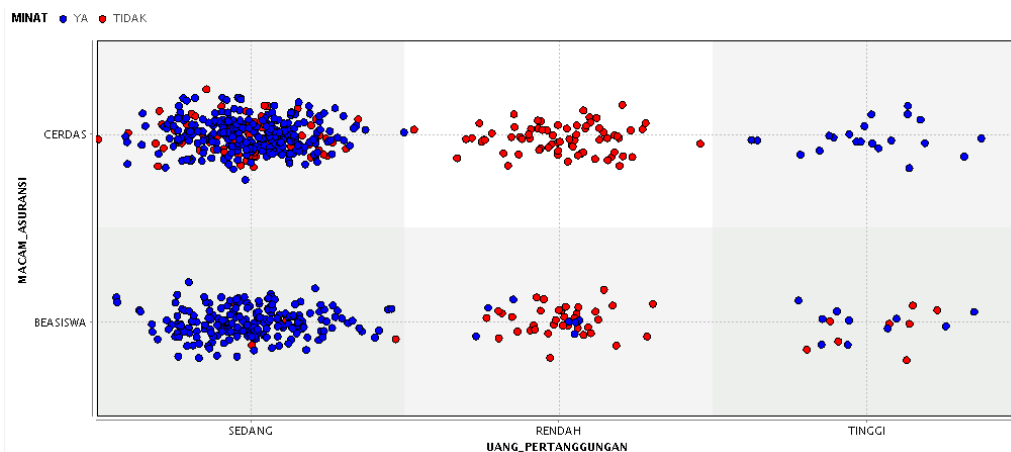


Gambar 3. Hasil proses *information gain* dan *gain ratio* pada *rapid miner 5*

Berdasarkan hasil pohon keputusan *information gain* dan *gain ratio* pada gambar 3 menghasilkan pohon keputusan yang sama, karena kedua metode tersebut termasuk jenis *data mining* klasifikasi *decision tree*. Atribut uang pertanggungan menempati posisi *root node* yang merupakan faktor paling berpengaruh dalam minat klien memilih produk asuransi pendidikan.

3.2 Implementasi *Rule Induction* pada *Rapid Miner 5*

Hasil proses *plot view rule induction* untuk memprediksi minat klien dalam memilih produk asuransi menggunakan aplikasi *rapid miner 5* ditunjukkan pada gambar 4.



Gambar 4. Hasil *plot view rule induction* pada *rapid miner 5*

Berdasarkan hasil *plot view rule induction* pada gambar 4, minat klien dengan uang pertanggungan “sedang”, “rendah”, maupun “tinggi” dengan macam asuransi “cerdas” maupun “beasiswa” menunjukkan bahwa minat klien “Ya” dalam memilih produk asuransi pendidikan lebih dominan pada uang pertanggungan “sedang” dan “tinggi” baik dari macam asuransi “beasiswa” maupun “cerdas”.

3.3 Implementasi *Information Gain*

3.3.1 Menentukan *root node* (simpul akar)

Tabel 2. Hasil *root node* dalam perhitungan *information gain*

Variabel	Atribut	Nilai <i>Information Gain</i>
X1	Macam asuransi	0,03226182
X2	Cara bayar	0,031610448
X3	Uang pertanggungan	0,24875197
X4	Premi	0,022171259

Dari hasil Tabel 2, uang pertanggungan menempati posisi *root node* karena memperoleh nilai *information gain* tertinggi, dapat dilihat pada gambar 5.



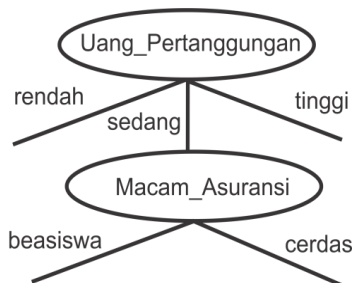
Gambar 5. *Root Node*

3.3.2 Menentukan *internal node* pertama

Tabel 3. Hasil *internal node* pertama dalam perhitungan *information gain*

Variabel	Atribut	Nilai <i>Information Gain</i>
X1	Macam asuransi	0,098810648
X2	Cara bayar	0,040070875
X4	Premi	0,019085654

Dari hasil Tabel 3, macam asuransi menempati posisi *internal node* pertama pada uang pertanggungan “sedang”, dapat dilihat pada gambar 6.



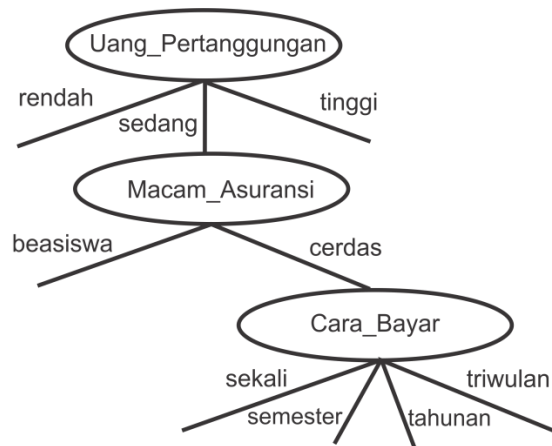
Gambar 6. *Internal node* pertama

3.3.3 Menentukan *internal node* kedua

Tabel 4. Hasil *internal node* kedua dalam perhitungan *information gain*

Variabel	Atribut	Nilai <i>Information Gain</i>
X2	Cara Bayar	0,090509426
X4	Premi	0,015179764

Dari hasil Tabel 4, cara bayar menempati posisi *internal node* kedua pada uang pertanggungan “sedang” dengan macam asuransi “cerdas”, dapat dilihat pada gambar 7.



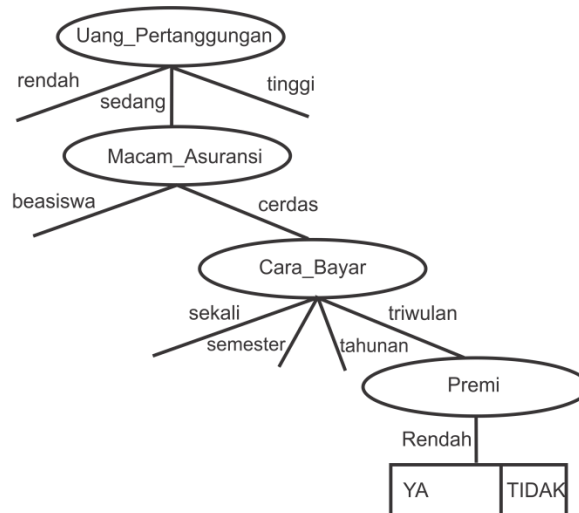
Gambar 7. *Internal node* kedua

3.3.4 Menentukan *internal node* ketiga dan *leaf node*

Tabel 5. Hasil *internal node* ketiga dan *leaf node* dalam perhitungan *information gain*

Variabel	Atribut	Nilai <i>Information Gain</i>
X4	Premi	0

Dari hasil Tabel 5, Premi menempati posisi *internal node* ketiga karena hasil *information gain* bernilai 0 jadi dapat disimpulkan bahwa uang pertanggungan “sedang” dengan macam asuransi “cerdas” cara bayar “tahunan” dan premi “rendah” menghasilkan *leaf node* minat “Ya”, dapat dilihat pada gambar 8.



Gambar 8. *Internal node* ketiga dan *leaf node*

3.4 Implementasi *Gain Ratio*

3.4.1 Menentukan *root node* (simpul akar)

Tabel 6. Hasil *root node* dalam perhitungan *gain ratio*

Variabel	Atribut	Nilai <i>Gain Ratio</i>
X1	Macam asuransi	0,033967112
X2	Cara_Bayar	0,033160159
X3	Uang pertanggungan	0,268661153
X4	Premi	0,095388594

Dari hasil Tabel 6, uang pertanggungan menempati posisi *root node* karena memperoleh nilai *gain ratio* tertinggi, dapat dilihat seperti gambar 9.



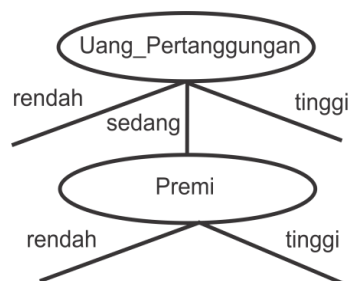
Gambar 9. *Root Node*

3.4.2 Menentukan *internal node* pertama

Tabel 7. Hasil *internal node* pertama dalam perhitungan *gain ratio*

Variabel	Atribut	Nilai <i>Gain Ratio</i>
X1	Macam_Asuransi	0,104917895
X2	Cara_Bayar	0,044849223
X4	Premi	0,160688495

Dari hasil Tabel 7, premi menempati posisi *internal node* pertama pada uang pertanggungan “sedang”, dapat dilihat pada gambar 10.



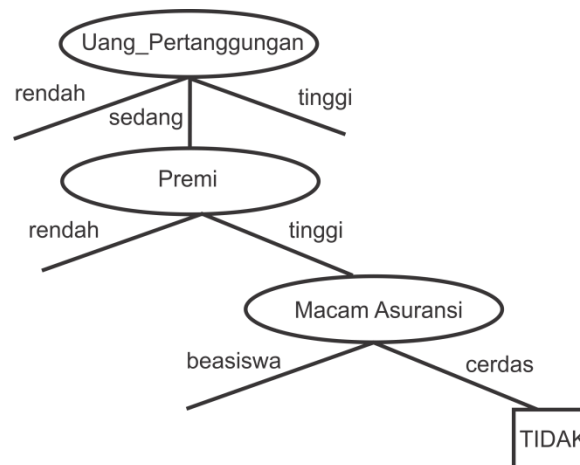
Gambar 10. *Internal node* pertama

3.4.3 Menentukan internal node kedua dan leaf node

Tabel 8. Hasil *internal node* kedua dan *leaf node* dalam perhitungan *gain ratio*

Variabel	Atribut	Nilai <i>Gain Ratio</i>
X1	Macam Asuransi	0,165531932
X2	Cara Bayar	0,02696039

Dari hasil Tabel 8, macam asuransi menempati posisi *internal node* ketiga karena hasil *gain ratio* tertinggi dan menghasilkan *leaf node* “Tidak” pada macam asuransi “cerdas” yang menghasilkan nilai *entropy* 0, dapat dilihat pada gambar 11.



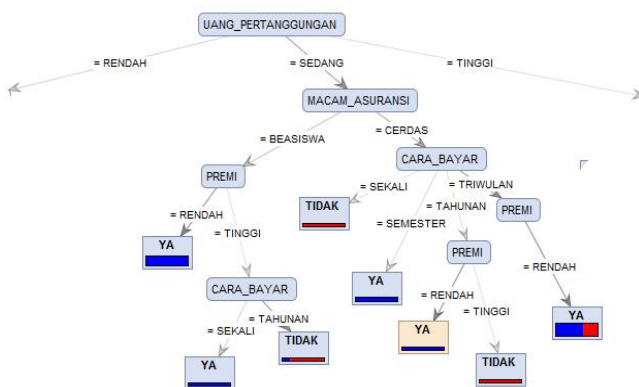
Gambar 11. *Internal node* kedua dan *leaf node*

3.5 Implementasi *Rule Induction*

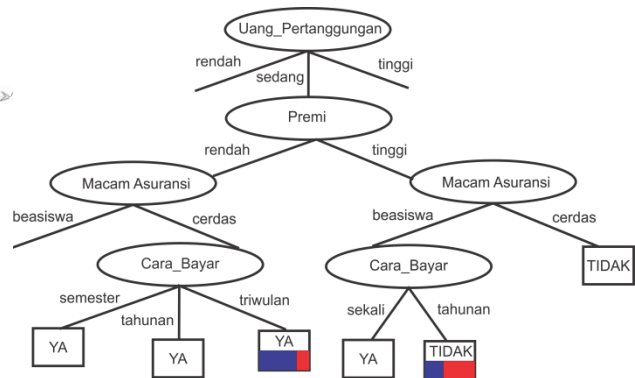
Berdasarkan hasil *plot view rule induction* seperti pada gambar 5, minat klien “Ya” lebih dominan pada uang pertanggungan “sedang” maupun “tinggi” baik dari macam asuransi “beasiswa” atau “cerdas” dalam penelitian minat klien memilih produk asuransi pendidikan.

3.6 Perbedaan Implementasi Hasil Implementasi

Perbedaan hasil *gain ratio* uang pertanggungan “sedang” yang dihasilkan dari analisa implementasi *rapid miner* 5 dengan implementasi perhitungan ditunjukkan pada gambar 12 dan 13.



Gambar 12. Hasil *rapid miner* 5



Gambar 13. Hasil perhitungan

Berdasarkan perbedaan hasil implementasi pada gambar 12 dan 13 menunjukkan posisi *internal node* pertama sampai posisi *leaf node* uang pertanggung “sedang” antara hasil *rapid miner 5* dengan hasil perhitungan tidak sama.

3.7 Hasil Komparasi Metode Data Mining

Musthafa (2015) menyatakan suatu model klasifikasi menggambarkan hasil benar dan salah disebut dengan *confusion matrix*, ditunjukkan pada Tabel 9.

Tabel 9. *Confusion Matrix*

<i>Predicted class</i>			
<i>Actual class</i>	Diterima	Diterima	Tidak Diterima
		<i>True Positive</i>	<i>False Negative</i>
		(TP)	(FN)
	Tidak Diterima		
		<i>False Positive</i>	<i>True Negative</i>
		(FP)	(TN)

Sehingga rumus *precision*, *recall*, dan *accuracy* terdapat pada **persamaan 8, 9, 10**.

Persamaan 8 Rumus *Precision* :

$$Precision = TP / (TP + FP) \times 100\% \quad (8)$$

Persamaan 9 Rumus *Recall* :

$$Recall = TP / (TP + FN) \times 100\% \quad (9)$$

Persamaan 10 Rumus *Accuracy* :

$$Accuracy = (TP + TN) / \text{Total Sampel} \times 100\% \quad (10)$$

Hasil komparasi *information gain*, *gain ratio*, dan *rule induction* dapat dilihat pada Tabel 10.

Tabel 10. Hasil komparasi

Komponen	<i>Information Gain</i>	<i>Gain Ratio</i>	<i>Rule Induction</i>
<i>Accuracy</i>	85,69 %	85,69 %	84,15 %
<i>Precision</i>	99,18 %	99,18 %	94,31 %
<i>Recall</i>	54,50 %	54,50 %	52,25 %

Berdasarkan hasil komparasi pada Tabel 10 dapat dilihat nilai *accuracy*, *precision*, dan *recall* pada *rule induction* lebih rendah daripada *information gain* dan *gain ratio*. *Information gain* dan *gain ratio* memiliki nilai yang sama, karena kedua metode ini termasuk dalam metode *data mining* klasifikasi *decision tree* sehingga metode ini adalah metode yang paling baik dalam penelitian.

3.8 Interpretasi *Information Gain* dan *Gain Ratio*

Dari hasil analisa didapatkan gambaran dalam menentukan minat klien memilih produk asuransi pada metode *information gain* dan *gain ratio* dengan menghasilkan pohon keputusan yang sama. Sehingga interpretasi yang terbentuk adalah :

1. Jika uang pertanggungan “rendah”, cara bayar “sekali” atau cara bayar “triwulan” akan menghasilkan minat “tidak”.
2. Jika uang pertanggungan “rendah”, cara bayar “semester” atau cara bayar “ tahunan” dengan macam asuransi “beasiswa” akan menghasilkan minat “ya”.
3. Jika uang pertanggungan “rendah”, cara bayar “semester” atau cara bayar “ tahunan” dengan macam asuransi “cerdas” akan menghasilkan minat “tidak”.
4. Jika uang pertanggungan “sedang”, macam asuransi “beasiswa” dengan premi “rendah” akan menghasilkan minat “ya”.
5. Jika uang pertanggungan “sedang”, macam asuransi “beasiswa” dengan premi “tinggi” dan cara bayar “sekali” akan menghasilkan minat “ya”.
6. Jika uang pertanggungan “sedang”, macam asuransi “beasiswa” dengan premi “tinggi” dan cara bayar “tahunan” akan menghasilkan minat “tidak”.
7. Jika uang pertanggungan “sedang”, macam asuransi “cerdas” dengan cara bayar “sekali” akan menghasilkan minat “tidak”.
8. Jika uang pertanggungan “sedang”, macam asuransi “cerdas” dengan cara bayar “semester” akan menghasilkan minat “ya”.
9. Jika uang pertanggungan “sedang”, macam asuransi “cerdas”, cara bayar “tahunan” maupun “triwulan” dengan premi “rendah” akan menghasilkan minat “ya”.
10. Jika uang pertanggungan “sedang”, macam asuransi “cerdas” dengan cara bayar “tahunan” dan premi “tinggi” akan menghasilkan minat “tidak”.
11. Jika uang pertanggungan “tinggi” dengan premi “rendah” akan menghasilkan minat “ya”.
12. Jika uang pertanggungan “tinggi”, premi “tinggi” dengan macam asuransi “beasiswa” akan menghasilkan minat “tidak”.
13. Jika uang pertanggungan “tinggi”, premi “tinggi” dengan macam asuransi “cerdas” akan menghasilkan minat “ya”.

3.9 Interpretasi *Rule Induction*

Interpretasi yang terbentuk pada metode *rule induction* berdasarkan gambar 4, adalah :

1. Uang pertanggungan “rendah”
 - a) Ya = Macam asuransi “beasiswa” merupakan produk asuransi yang paling diminati.
 - b) Tidak = Macam asuransi “cerdas” merupakan produk asuransi kurang diminati.

2. Uang pertanggungan “sedang”
 - a) Ya = Macam asuransi “beasiswa” maupun macam asuransi “cerdas” merupakan produk asuransi yang paling diminati.
3. Uang pertanggungan “tinggi”
 - a) Ya = Macam asuransi “cerdas” merupakan produk asuransi yang paling diminati.
 - b) Tidak = Macam asuransi “beasiswa” merupakan produk asuransi yang kurang diminati.

4. PENUTUP

Berdasarkan dari analisa hasil penelitian ini, maka dapat disimpulkan bahwa :

1. Uang pertanggungan menempati posisi sebagai *root node* (simpul akar) dalam metode *information gain* dan *gain ratio* yang menjadikan variabel uang pertanggungan adalah variabel paling berpengaruh pada penelitian minat klien dalam memilih produk asuransi pendidikan.
2. Berdasarkan hasil *plot view rule induction* minat klien “Ya” lebih dominan pada uang pertanggungan “sedang” maupun “tinggi” baik dari macam asuransi “beasiswa” atau “cerdas”.
3. Berdasarkan nilai *accuracy*, *precision*, dan *recall* pada metode *information gain* dan *gain ratio* memiliki nilai yang sama yaitu nilai *accuracy* sebesar 85,69 %, nilai *precision* sebesar 99,18 %, dan nilai *recall* sebesar 54,50 %. Sedangkan metode *rule induction* menghasilkan nilai *accuracy* sebesar 84,15 %, nilai *precision* sebesar 94,31 %, dan nilai *recall* sebesar 52,25 % . Sehingga metode *information gain* dan *gain ratio* dalam penelitian ini menjadi metode yang sama unggul.
4. Adanya perbedaan hasil implementasi *rapid miner* 5 dengan hasil perhitungan pada metode *gain ratio* uang pertanggungan “sedang” yang terletak pada posisi *internal node* pertama sampai *leaf node*.

Adapun rekomendasi peneliti untuk perusahaan, klien dan peneliti selanjutnya dari hasil penelitian ini, adalah sebagai berikut :

1. Perusahaan dalam menawarkan produk asuransi pendidikan lebih memperhatikan jumlah uang pertanggungan dan premi yang dibayarkan agar dapat sesuai dengan harapan klien.
2. Meningkatkan mutu produk asuransi pendidikan produk mitra beasiswa agar minat klien bertambah dan tetap menjaga mutu produk mitra cerdas agar produk ini tetap diminati.
3. Produk yang disarankan untuk nasabah yang hendak mulai asuransi adalah produk pendidikan mitra cerdas karena produk tersebut menghasilkan investasi yang meningkat.
4. Penelitian ini menggunakan *rapid miner* versi 5, untuk penelitian selanjutnya disarankan menggunakan *rapid miner* versi terbaru yang lebih *update* dengan kelengkapan fitur.

5. Tidak adanya aplikasi *website* dalam penelitian ini, sehingga perlu ditambahkan aplikasi *website* bagi peneliti selanjutnya.

PERSANTUNAN

Terima kasih untuk dosen pembimbing ibu Umi Fadlilah, S.T, M.Eng, kepala administrasi bumiputera 1912 karanganyar bapak Drs. Sudadi, karyawan bumiputera 1912 Karanganyar Devis Raizal Septiawan S.Kom, dan rekan-rekan yang telah membantu dalam penelitian dan penulisan naskah publikasi.

DAFTAR PUSTAKA

- Anita, S. (2015). Data Mining Memprediksi Masyarakat terhadap Asuransi Jiwa dengan Metode Interpolasi Linier. *Jurnal Riset Komputer*, 2, 62-65.
- Kumbhar, P., & Mali, M. (2016). A Survey on Feature Selection Techniques and Classification Algoritms for Efficient Text Classification. *International Journal of Science and Research*, 5, 1267-1275.
- Mahanani, L. (2016). Perbandingan 3 Metode Dalam Data Mining untuk Penentuan Kadar Kolesterol di RSUD Dr. Moewardi Surakarta. Surakarta: Universitas Muhammadiyah Surakarta.
- Musthafa, A., Suyono, H., & Sarosa, M. (2015). Perbandingan Kinerja Algoritma C.45 dan AHP-TOPSIS Sebagai Pendukung Keputusan Proses Seleksi Penerima Beasiswa. *Jurnal EECCIS*, 9, 109-114.
- Nugroho, Y. S. (2016). Modul Praktikum Data Warehousing dan Data Mining. Surakarta: Universitas Muhammadiyah Surakarta.
- Ochtario, R., & Karnila, S., (2013). Sistem Data Mining untuk Mengetahui Tingkat Kecenderungan Memilih Menu Makanan dengan Metode Association Rule Mining. *Jurnal Informatika*, 13, 80-89.
- Pratiwi, L.M., Diana., & Agustin E,P. (2017). Penerapan K-means Clustering untuk Memprediksi Minat Nasabah pada PT. Asuransi Jiwa Bersama 1912 Bumiputera Prabumulih.
- Usman, T. (2016). Perbandingan 3 Metode Dalam Data Mining untuk Penentuan Hipertensi di Rumah Sakit Umum Daerah Dr. Moewardi Surakarta. Surakarta: Universitas Muhammadiyah Surakarta.